Discussion

Steven G. Rivkin

The passage of the No Child Left Behind Act (NCLB) represents a major shift in the federal role in elementary and secondary education. Whether it succeeds in raising academic achievement and improving subsequent economic and social outcomes remains to be seen. As Hanushek and Raymond point out, the recent, nationwide implementation of NCLB makes it very difficult to identify the impacts of the reform. This difficulty led them to use existing evidence to predict the likely effects of the law. A sensible approach. Of course, the predictive power of the existing evidence depends on both the strength of that evidence and its relevance to the accountability reforms mandated by NCLB.

The paper begins by documenting the growing body of evidence showing that mathematics achievement raises earnings and economic growth, both directly and indirectly, by increasing educational attainment. The evidence indicates that school quality improvements could have substantial economic benefits and affect both average earnings and the degree of inequality. Therefore education reforms could have a large economic payoff.

Questions do remain about the overall magnitude of the effects and the impact of school quality improvements on the earnings distribution. If those who have a higher expected return to learning mathematics tend to receive better instruction or put forth higher effort, returns based on earnings differences may understate the value of mathematics knowledge for some (engineers, for example) and overstate the value for others (writers or laborers, for example). Higher mathematics achievement may also be related to other determinants of earnings, including family income and wealth, in which case the observed return to mathematics knowledge could conflate the actual return and related advantages of growing up in a higher-income family. Alternatively, evidence on the returns to mathematics skills, such as the returns depicted in Figure 1, may understate the true return by assuming that the increase in one's earnings is constant in percentage terms throughout one's career. Most studies use samples of younger workers early in their careers, and those with higher mathematics skills may be "investing" more in the sense of foregoing current earnings to obtain additional skills that will increase future earnings. Yet despite these and other concerns, the weight of the evidence points to a strong economic payoff to raising mathematics knowledge, meaning that the benefits of accountability reforms depend on their effects on school quality.

Accountability systems adopted prior to NCLB facilitate the study of accountability effects. The authors point out the substantial variation in program structure and emphasize the importance of design details, including the type and severity of sanctions, generosity of rewards, and method for estimating school and teacher effects. Long-run success of accountability systems depends on a consensus of belief that teacher and school evaluations and rankings are based on actual quality and not on the skills that students bring to the classroom. Proper assessment requires the development of comprehensive data systems that can also be used in the study of state education systems more broadly. Not only would flawed assessment fail to

Steven G. Rivkin is an associate professor of economics at Amherst College.

Federal Reserve Bank of St. Louis *Regional Economic Development*, 2006, *2*(1), pp. 62-63.

reward or sanction based on true quality, it would discourage high-quality teachers and administrators from remaining in and entering public education.

Hanushek and Raymond argue that variation in the timing and structure of state accountability systems adopted prior to the passage of NCLB can be used to estimate the effects of reforms that are similar in many respects to NCLB. Complications arise because differences in the timing and extent of reforms may be related to other factors that affect achievement, making it difficult to isolate the accountability effects. For example, states may implement a number of reforms in response to low performance or changes in the political climate. These include alignment of the curriculum more closely to the type of material covered on the National Assessment of Educational Progress (NAEP) examination, the test used in this and other papers. Consequently, observed changes in test scores could result from a number of factors, necessitating the identification of a comparison group that would provide a benchmark of achievement changes not resulting directly from the newly adopted accountability system. States yet to adopt accountability systems are good candidates, but they may fail to capture other systematic changes experienced by the accountability adopters.

The results suggest that high-stakes accountability systems raise achievement, but there is reason to interpret the findings with some caution due to the variation in program characteristics—the possibility that other factors are contaminating the results—and to the pattern of the results. Specifically, given that state tests tended to focus on less-advanced skills, improvements in schools serving poor and minority students with lower initial scores would be more likely to be captured on the tests. Moreover, one might expect the introduction of high-stakes examinations to have a larger impact on schools serving predominantly minority and lower-income students whose families likely placed less pressure on school administrators to raise quality than did families of middle and upper class students who can more easily opt out of underperforming schools by switching to the private sector or moving to a different community.

Additional concerns revolve around the question of the size of the effects on longer-run outcomes, including academic attainment, earnings, and economic growth. The fact that the NAEP data is not the assessment used in the accountability programs certainly mitigates the likelihood that test score increases do not carry over to longer-run outcomes or even other tests in the same subjects. Nonetheless, reports of extensive time devoted to test-taking instruction and additional emphasis on the test material vis à vis other subjects highlights the importance of structuring the incentives correctly.

Finally, although frustration with resource-based policies has been a catalyst for the growing demand for accountability and other incentive programs, expanded resources and more-rigorous incentives are not mutually exclusive tools for improving the schools. In fact, one measure of the success of accountability efforts is the extent to which these reforms increase the return to financial investments in the public schools.